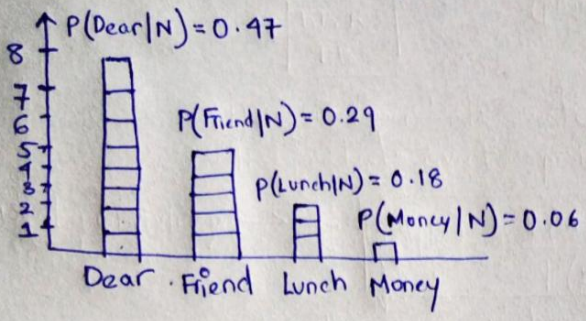
NAIVE BAYES (BY STATQUEST)

Example → Imagine we received normal message from friends and family. And we also received spam (unwanted message that are usually scams or unsolicated ad)

AIM → To filter out the spam messages.

So first thing we do is **make histogram** of all words that occurs in the normal message from family and friends. We can use **histogram to calculate the probability** of seeing each word in **normal message**.



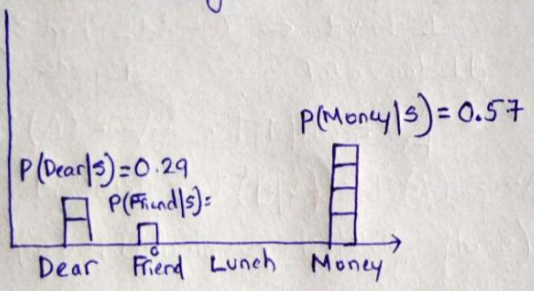
$$P(\text{Dear}|\text{Normal}) = \frac{8}{17} = \frac{\text{"Dear" word}}{\text{Total word in Normal messages}} = 0.47$$

$$P(\text{Friend}|\text{Normal}) = \frac{5}{17} = 0.29$$

$$P(\text{Lunch}|\text{Normal}) = \frac{3}{17} = 0.18$$

$$P(\text{Money}|\text{Normal}) = \frac{1}{17} = 0.06$$

→ Now **plot histogram** of all word that **occurs in the spam**.



$$P(\text{Dear}|\text{Spam}) = \frac{2}{7} = 0.29$$

$$P(\text{Friend}|\text{Spam}) = \frac{1}{7} = 0.14$$

$$P(\text{Lunch}|\text{Spam}) = \frac{0}{7} = 0$$

$$P(\text{Money}|\text{Spam}) = \frac{4}{7} = 0.57$$

→ Now, since 8 of the 12 messages are normal messages, our **initial guess is 0.67**

$$P(N) = \frac{8}{8+4} = 0.67$$, this is also called a **Prior Probability**.

Suppose in a message we get a message "Dear Friend". So the probability, the message is Normal = $P(\text{Normal}) \times P(\text{Dear}|\text{Normal}) \times P(\text{Friend}|\text{Normal})$

$= 0.67 \times 0.47 \times 0.29 = 0.09$

so we got 0.09 as a score that Dear Friend gets if its a Normal Message. So the eqn is

$P(\text{Normal}|\text{Dear Friend}) = 0.09$, is proportion to the probability that the message is normal given that it say Dear Friend.

Suppose now we say, message is spam.

$$P(\text{Spam}) = \frac{4}{4+8} = 0.33$$

"Dear Friend" and we consider it as Spam $= P(S) * P(\text{Dear}|S) \times P(\text{Friend}|S)$

$= 0.33 \times 0.29 \times 0.14 = 0.01$

$P(\text{Spam}|\text{Dear Friend}) = 0.01$

So, $P(\text{Normal}|\text{Dear Friend}), 0.09 > P(\text{Spam}|\text{Dear Friend}), 0.01$.

So we will decide that Dear Friend is a Normal message.

Therefore it can be **used in classification algorithm**, termed 2 class spam & not spam.

$P(\text{Normal}|\text{Dear Friend}) = P(\text{Normal}|\text{Friend Dear})$, **it ignore relationship between word so naive**

By ignoring among words, **Naive bayes have high Bias** and have **low variance**.